Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 4907–4910, June 1991
Genetics

Evidence for retrotransposition of the *I* factor, a *LINE* element of *Drosophila melanogaster*

(transposition/reverse transcription/intron/hybrid dysgenesis)

ALAIN PÉLISSON*†, DAVID J. FINNEGAN‡, AND ALAIN BUCHETON§¶

*Laboratoire de Génétique, Université Blaise Pascal, 63177 Aubière cedex, France; ‡Institute of Cell and Molecular Biology, University of Edinburgh, King's Buildings, Edinburgh EH9 3JR, Scotland; and §Centre de Génétique Moléculaire, Centre National de la Recherche Scientifique, 91198 Gif-sur-Yvette, France

Communicated by Walter J. Gehring, March 18, 1991 (received for review November 9, 1990)

ABSTRACT **_LINE_s are transposable elements found in various eukaryotes such as plants, protists, insects, and mammals. Their transposition is usually difficult to study, particularly in humans, where some diseases have been shown to result from _LINE_ insertion mutations. This is due to the fact that most copies of any particular family of elements are defective and that their transposition frequency is low. By contrast, the _I_ factor of _Drosophila melanogaster_ transposes at high frequency during _I–R_ hybrid dysgenesis and is a good model for studying the _LINE_ element superfamily. _LINE_s encode putative polypeptides showing similarities with viral reverse transcriptases but, unlike viral retrotransposons, they do not have terminal repeats and their ability to transpose by reverse transcription has previously only been inferred from structural analysis. Here we present direct evidence for _LINE_ retrotransposition. Transposition of an _I_ factor marked by an intron resulted in accurate removal of the intron.**

The *I* factor is a transposable element of *Drosophila melanogaster* related to *LINE*s. These elements are found in a variety of phyla including plants, protists, insects, and mammals (1). Particularly significant is their quantitative and qualitative contribution to mammalian genome organization and expression. For instance, in humans up to 5% of the genome consists of the same *LINE* family (*L1*). *L1* elements have been shown to transpose, and *de novo* insertions of these elements have been found associated with hemophilia A and breast carcinoma (2, 3).

*LINE*s are devoid of terminal repeats but have an A-rich sequence, usually preceded by a polyadenylylation signal, at their 3′ ends. They contain two long open reading frames, one of which encodes a putative polypeptide showing similarities with viral reverse transcriptases (1). They are thought to transpose by reverse transcription of an RNA intermediate, but this is based only on structural studies (4–6). Analysis of the mechanism of *LINE* element transposition is particularly difficult in mammals because a high proportion of these elements are defective and the frequency of transposition is very low.

I factors, which are responsible for the *I–R* system of hybrid dysgenesis in *D. melanogaster* (7, 8), have properties that overcome these difficulties. *I–R* hybrid dysgenesis is a syndrome that includes female infertility and increased frequencies of mutations and is associated with frequent replicative transposition of *I* factors. There are two categories of strains in *D. melanogaster* with respect to the *I–R* system: inducer strains that contain active *I* factors and reactive strains that do not. Crosses between inducer males and reactive females produce dysgenic F_1 females that are poorly fertile. Mutations occur at high frequency in their germ line. *I* factors are stable in inducer strains but are activated and transpose at high frequency in the germ line of dysgenic females (9). They never transpose in males.

Active *I* factors have been cloned. They are 5.4 kilobases (kb) long and are devoid of terminal repeats. They terminate at the 3′ end by several repeats of the sequence TAA. They contain two long open reading frames. One of them, ORF2, encodes a putative polypeptide showing similarities with reverse transcriptases and RNase H (10).

P–M dysgenesis is an independent system of hybrid dysgenesis that has some characteristics in common with *I–R* dysgenesis but involves transposable elements, *P* elements, that use a strictly DNA-dependent transposition mechanism (11). Strains completely devoid of *P* elements are called *M* strains, whereas strains containing active *P* elements are called *P* strains.

Transposition by reverse transcription of an RNA intermediate has been clearly demonstrated for *Ty1* elements in *Saccharomyces cerevisiae* by showing that an intron inserted within the element is removed during transposition (12). We have used a similar method to investigate the mechanism of transposition of the *I* factor of *D. melanogaster*, a member of a class of elements that are structurally quite distinct from retrovirus-like elements such as *Ty1*. For this, we have introduced an intron into an *I* factor. Its transposition resulted in precise removal of the intron, providing direct evidence for retrotransposition of *LINE*s. Our results also indicate that trans-complementation of elements of this type is possible.

MATERIALS AND METHODS

Construction of the *πI* Element. A marked *I* factor (*πI*) was constructed by introducing an intron-containing fragment of the *P* element of *D. melanogaster* into the functional *I* factor of clone pI407 (13, 14) (see Fig. 1). For this, the *Pvu* II/*Pvu* II restriction fragment of clone p*π*25.7 (15), which contains the second intron of the *P* factor, was cloned in the same transcriptional orientation as *I* factor sequences in *Kpn* I-cleaved pI407 (8) after treatment with T4 DNA polymerase. The 7.1-kb *Sal* I/*Sal* I fragment containing this marked *I* element, *πI*, was cloned into the *Sal* I site of the pUChsneo transformation vector (16), giving clone pPI.

Construction of Transgenic Lines. Clone pPI was microinjected into embryos (G_0) of the reactive strain Cha together with the p*π*25.7 wing-clipped helper plasmid as described (17). Transgenic flies were selected among the progeny of these embryos by growing G_1 larvae on food containing G418 (1 mg/ml). We recovered one transformed line in which the *πI* element containing *P*-element transposon was located on the X chromosome. It was called *π*IX. This insertion was subsequently moved three times to chromosome 3 of the inducer strain DcxF by *P*-mediated transposition using the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

†Present address: Department of Biology, The Johns Hopkins University, Baltimore, MD 21218.
¶To whom reprint requests should be addressed.

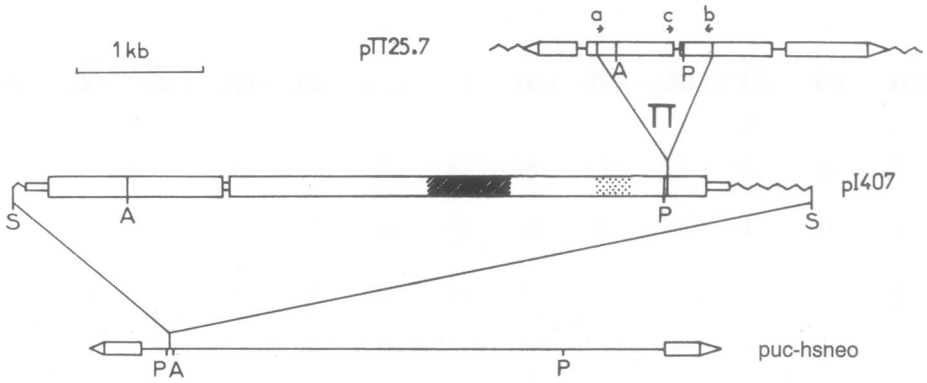

FIG. 1. Map of the *πI* element. The *P* and *I* elements in clones pπ25.7 (15) and pI407 (13) are drawn in the conventional 5′ to 3′ orientation. They are both flanked by genomic DNA represented by wavy lines. The terminal noncoding sequences and the exons and introns of *P* are represented by open arrowheads, open boxes, and thin lines, respectively. The orientations and positions of the primers used to amplify (a and b) and sequence (c) the *π* fragment are shown by arrows. The boxes in the *I* factor correspond to ORF1 and ORF2 and the hatched and stippled regions represent amino acid sequences similar to reverse transcriptase and RNase H, respectively. The pUChsneo transformation vector (16) contains pUC8 sequences and the hsneo selectable marker flanked by *P* factor terminal repeats. A, *Ava* I; P, *Pst* I; S, *Sal* I.

P(ry$^+$ Δ2-3)(99B) chromosome as a donor of transposase (11). The corresponding transgenic strains were called *π*ID1, *π*ID2, and *π*ID3.

DNA Amplification Using PCR. Genomic DNAs were extracted according to a proteinase K/phenol/RNase/phenol procedure. The *Pst* I fragment of *πI* was amplified by two successive steps of PCR (18). One microgram of genomic DNA was mixed with 0.25 μg each of oligonucleotides a and b (see Fig. 1) and 0.5 μl (2.5 units) of *Taq* polymerase (Perkins–Elmer/Cetus) in 100 μl of PCR buffer [6.7 mM $MgCl_2$/16.6 mM $(NH_4)_2SO_4$/67 mM Tris·HCl, pH 8.8/0.25 mM each dNTP]. The reaction mixture was overlaid with mineral oil and subjected to 30 temperature cycles (94°C denaturing, 1 min; 60°C annealing, 1 min; 72°C extension, 1.5 min). About 2 pg (10^{-5} of the PCR product) was amplified again in the same conditions as described above for sequencing.

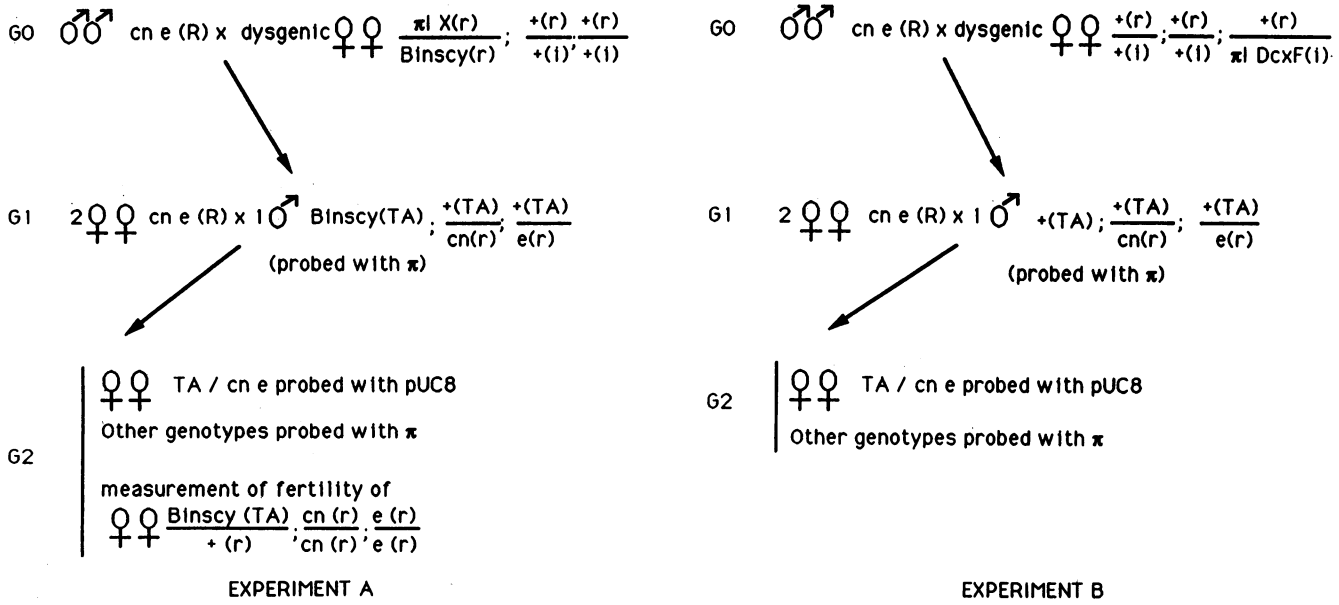

FIG. 2. Mating scheme to detect *I–R* dysgenesis-dependent transposition of the *πI* element from various insertion sites. Only *M* strains were used so that no transposition events could result from *P*-element activity. Chromosomes originating from an inducer or a reactive origin are symbolized by (i) and (r), respectively. (TA) is a target chromosome that might have acquired *I* and *πI* elements in the germ line of the G_0 dysgenic females. Experiment A was designed to detect transposed copies of the *πI* element located on the X chromosome of the *π*IX stock and experiment B was designed to detect transposed copies of the *πI* element located on the DcxF chromosome of stocks *π*ID1, *π*ID2, and *π*ID3. In experiment A, the G_0 dysgenic females were obtained by mating *π*IX reactive females with males heterozygous for inducer and reactive chromosomes providing the Binscy reactive X chromosome. In experiment B, the dysgenic females were obtained by crossing females of the wild-type reactive strain Cha and inducer males providing one or another of the three DcxF chromosomes containing the *πI* element. The Binscy and DcxF chromosomes contain crossing-over inhibitors to prevent recombination between *πI* carrying chromosomes and the homologous chromosomes of the target genome, TA. The TA genome was transmitted to G_1 males that were mated individually with *cn;e* reactive females (*cn* and *e* are recessive markers of the second and third chromosomes, respectively). They were then hybridized with the *π* probe according to the squash blot procedure (20) to select individuals carrying sequences from the *πI* element. To eliminate translocations of the whole pUChsneo *πI* vector as a result of exceptional recombination, the wild-type female progeny of positive G_1 males were probed with pUC8. Any line having homology to this probe was discarded. Lines positive with the *π* probe and negative with pUC8 contained transposed copies of the marked *πI* elements. The chromosomal locations of these transposed *πI* elements were determined by probing flies of various G_2 genotypes with *π*. Two transposed copies were recovered in one G_1 male from *π*ID1 in experiment B. One was located on the X chromosome (*π*ITX) and the other was on chromosome 3 (*π*IT3).

Sequencing Reactions. The Sequenase kit (United States Biochemical) was used to determine the nucleotide sequence as follows. One-third of the microconcentrator-purified PCR product obtained as described above was denatured in 0.2 M NaOH and annealed with 50 pM primer (see Fig. 1, oligonucleotide c). Synthesis was performed with 3 pM [α-^{32}P]dATP (3000 Ci/mmol; 1 Ci = 37 GBq) and the labeling mix was diluted 10 times. These experiments, including genomic DNA extraction and PCR amplification, were repeated twice for each DNA. Identical results were obtained in each case.

Other Procedures. All other procedures were carried out as described (19).

RESULTS AND DISCUSSION

The πI Element Cannot Transpose Autonomously. We have constructed an element, πI, to test whether *I* factors transpose via an RNA intermediate. This was made by inserting an intron-containing fragment into the second open reading frame of an active *I* factor (13, 14). An 896-base-pair fragment (π) containing the second intron of the *P* factor was used (Fig. 1). The marked πI element was introduced into the germ line of a reactive strain by a standard *P*-mediated transformation technique (16, 17). Four transgenic lines were obtained from this experiment. In one of them (πIX), the marked πI element was inserted on the X chromosome. In the other three (πID1, πID2, and πID3) it was inserted on a third chromosome carrying the DcxF rearrangement. The πI elements in the transgenic lines did not show visible rearrangements (see Fig. 3, lanes X, D1, D2, and D3).

The marked *I* elements are unable to transpose in these strains (data not shown) because of the mutation caused by the insertion of *P* sequences into a coding region. They were therefore transferred to dysgenic females by appropriate crosses so that they could be complemented by transposing *I* factors (Fig. 2).

The Intron Is Accurately Spliced During Transposition of the πI Element. For each transgenic line, ≈150 G_1 males were individually screened for homology to the π probe in experiments A and B (Fig. 2). From the πI element of line πID1, two bona fide transposition events were isolated in the progeny of one G_1 male according to the procedure described in Fig. 2. The chromosomes carrying the transposed πI elements in this G_1 male were identified by screening its G_2 progeny for π sequences in squash blot experiments (20). Females of the genotype *TA*/+;*cn*/*cn*;*e*/*e* and males of the genotype +/*Y*;*cn*/*cn*;*TA*/*e* were positive while all +/*Y*;*TA*/*cn*;*e*/*e* males were negative. This indicates that one transposed copy of the marked πI element is inserted on the X chromosome while the other is inserted on chromosome 3.

Restriction analysis of these chromosomes, called πITX and πIT3, is shown in Fig. 3 (lanes TX and T3). Sequences hybridizing to the π probe are on *Sal* I fragments that are larger than the 7.1-kb fragment of the donor (Fig. 3*a*), whereas the 4.3-kb internal *Ava* I fragment of the πI factor has been conserved (Fig. 3*b*). These results confirm that the πI element transposed independently of flanking sequences from the *P* transformation vector. This appears to have happened twice in the same germ cell of a dysgenic female (see Fig. 2). The *Pst* I fragments that contain the intron in the πI donor element are ≈60 bp shorter in πITX and πIT3 than in the πI element (Fig. 3*c*). The sequence of the appropriate region of these fragments indicates that this difference results from precise removal of the intron in each case (Fig. 4).

Therefore, the πI element must have transposed via an RNA intermediate that was reverse transcribed after being processed. Reverse transcription is presumably the mechanism by which other *LINE* elements transpose. A full-length *I* factor RNA has been recently identified in the ovaries of dysgenic females (21). It could be the intermediate required for transposition.

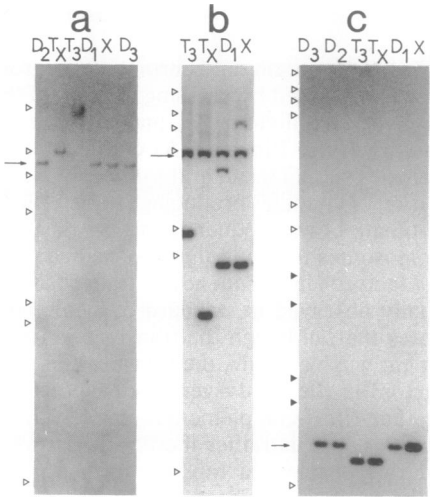

FIG. 3. Restriction analysis of various πI insertions. The following restriction enzymes were used: *Sal* I (*a*), *Ava* I (*b*), and *Pst* I (*c*). X is the πIX stock containing the original πI transgene on the X chromosome. D1, D2, and D3 indicate stocks with the three πI-carrying DcxF chromosomes. TX and T3 are the πITX and πIT3 stocks isolated from experiment B in Fig. 2. The positions of λ 23.6-, 9.6-, 6.7-, 4.3-, 2.3-, 2.0-, and 0.56-kb *Hin*dIII (open arrowheads) and λ 1.57-, 1.32-, 0.93-, and 0.84-kb *Hin*dIII/*Eco*RI fragments (solid arrowheads) are indicated. Arrows indicate genomic fragments corresponding in size to the control fragments of the pPI construct. Genomic DNA (3–4 μg) was digested with various restriction enzymes, electrophoresed on 0.7% (*a*), 1% (*b*), and 1.3% (*c*) agarose gels, transferred to Amersham Hybond N filters, and hybridized with the *Pst* I/*Pst* I fragment covering most of the π marker (see Fig. 1). Hybridization, washing, and autoradiography followed the suppliers' procedures. In *b*, a common band corresponding to the *Ava* I fragment internal to the πI element (see Fig. 1) is observed in all lines. The other major band corresponds to the fragment included between the *Ava* I site of the *P* marker inserted within πI and the *Ava* I site to the right of πI in Fig. 1, which is either in the pUChsneo vector (X and D1) or in the target sequences flanking πI (T3 and TX), confirming that in πIT3 and πITX, the πI element transposed independently of pUChsneo sequences.

Trans-Complementation of *I* Elements Can Occur. From the experiments reported above, it appears that frequency of transposition of the πI element in the G_0 dysgenic females (Fig. 2) was $<10^{-2}$. This is very low compared to the transposition frequency of active *I* factors usually observed in dysgenic females. In experiment A (Fig. 2), this frequency

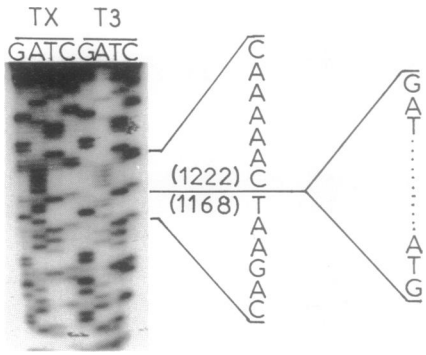

FIG. 4. Nucleotide sequence of the π fragments after transposition of the πI element. TX and T3 correspond to πITX and πIT3 stocks isolated from experiment B (Fig. 2). The nucleotide sequence at the breakpoint of the deletion is given with coordinates of the *P* factor sequence (15); 1168 and 1222 are the end of the second exon and the beginning of the third exon, respectively.

was estimated by measuring the fertility of 11 samples of G_2 females of the genotype Binscy (*TA*)/ + (*r*);*cn* (*r*)/*cn* (*r*);*e*(*r*)/ *e*(*r*). Such females contain only chromosomes from reactive origin and were obtained by crossing females of the reactive strain *cn*;*e* with 11 individual G_1 males. Acquisition of an active *I* factor by the Binscy (TA) chromosome in the G_0 dysgenic females should result in sterility of these G_2 females. The proportion of G_1 males producing such sterile G_2 females gives an estimate of the frequency of transposition. Four of the 11 chromosomes tested in this way had acquired at least one active *I* factor. This frequency of transposition is similar to that usually observed in standard dysgenic females (22). This indicates that, although the transposition frequency of the *πI* element was very low, the transposition frequency of active *I* factors in the G_0 dysgenic females was very high, suggesting that trans-complementation of transposition of elements of this type is rather inefficient. However, it can occur, and this provides a tool for further more detailed studies of the mechanism of transposition.

This work was initiated in the laboratory of D.J.F. where A.P. stayed with the support of a short-term European Molecular Biology Organization fellowship. We thank V. Corces for reading the manuscript and for sharing his facilities at the end of the work. This work has been financed by the Centre National de la Recherche Scientifique, Association pour la Recherche sur le Cancer, Fondation pour la Recherche Médicale, Ligue Nationale contre le Cancer, and Medical Research Council.

1. Hutchinson, C. A., III, Hardies, S. C., Loeb, D. D., Shehee, W. R. & Edgell, M. H. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington, DC), pp. 593–617.
2. Kazazian, H. H., Jr., Wong, C., Youssufian, H., Scott, A. F., Phillips, D. G. & Antonarakis, S. E. (1988) *Nature (London)* **332**, 164–166.
3. Morse, B., Rotherg, P. G., South, V. J., Spandorfer, J. M. & Astrin, S. M. (1988) *Nature (London)* **333**, 87–90.
4. Rogers, J. (1985) *Int. Rev. Cytol.* **93**, 187–280.
5. Schwartz-Sommer, Z., Leclerq, L., Göbel, E. & Saedler, H. (1987) *EMBO J.* **6**, 3873–3880.
6. Busseau, I., Pélisson, A. & Bucheton, A. (1989) *Nucleic Acids Res.* **17**, 6939–6945.
7. Bucheton, A. (1990) *Trends Genet.* **6**, 16–21.
8. Finnegan, D. J. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington, DC), pp. 503–518.
9. Picard, G. (1979) *Genetics* **91**, 455–471.
10. Fawcett, D. H., Lister, C. K., Kellett, E. & Finnegan, D. J. (1986) *Cell* **47**, 1007–1015.
11. Engels, W. R. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington, DC), pp. 437–484.
12. Boeke, J. D., Garfinkel, D. J., Styles, C. A. & Fink, G. R. (1985) *Cell* **40**, 491–500.
13. Bucheton, A., Paro, R., Sang, H. M., Pélisson, A. & Finnegan, D. J. (1984) *Cell* **38**, 153–163.
14. Pritchard, M., Dura, J. M., Pélisson, A., Bucheton, A. & Finnegan, D. J. (1988) *Mol. Gen. Genet.* **214**, 533–540.
15. O'Hare, K. & Rubin, G. M. (1983) *Cell* **34**, 25–35.
16. Steller, H. & Pirrotta, V. (1985) *EMBO J.* **4**, 167–171.
17. Spradling, A. C. (1986) in *Drosophila: A Practical Approach*, ed. Roberts, D. B. (IRL, Oxford), pp. 175–197.
18. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239**, 487–491.
19. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
20. Anxolabehere, D., Nouaud, D., Périquet, G. & Tchen, P. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5418–5422.
21. Chaboissier, M. C., Busseau, I., Prosser, J., Finnegan, D. J. & Bucheton, A. (1990) *EMBO J.* **9**, 3557–3563.
22. Picard, G. (1978) *Mol. Gen. Genet.* **164**, 235–247.